

Theory of single-electron heat engines coupled to electromagnetic environments

Tomi Ruokola¹ and Teemu Ojanen^{2,3}

¹ *Department of Applied Physics, Aalto University, P.O. Box 11100, FI-00076 Aalto, Finland*

² *Low Temperature Laboratory, Aalto University, P.O. Box 15100, FI-00076 Aalto, Finland and*

³ *Physics Department, Harvard University, Cambridge, Massachusetts 02138, USA*

(Dated: February 20, 2012)

We introduce a new class of mesoscopic heat engines consisting of a tunnel junction coupled to a linear thermal bath. Work is produced by transporting electrons up against a voltage bias like in ordinary thermoelectrics but heat is transferred by microwave photons, allowing the heat bath to be widely separated from the electron system. A simple and generic formalism capable of treating a variety of different types of junctions and environments is presented. We identify the systems and conditions required for maximal efficiency and maximal power. Very high figures of merit and efficiencies above the Curzon–Ahlborn limit are achievable.

PACS numbers: 73.23.Hk, 73.50.Lw, 72.70.+m, 44.40.+a

I. INTRODUCTION

Conversion of heat to work, and, in particular, recovery of waste heat produced by electronic components is a problem of great and ever increasing importance. Solid-state thermoelectric systems¹ are ideally suited for this purpose since they are easily integrated with the rest of the circuitry on the microchip. Mesoscopic heat transfer devices are a promising class of thermoelectric systems due to their ease fabrication, control, and measurement,² and because sharp features in the energy spectrum, a requirement for efficient operation,³ are readily available. Due to their small size they can also be used to study foundational issues, such as the importance of fluctuations⁴ and the fundamental limits of heat engine performance.⁵

Arguably the simplest mesoscopic heat engine consists of a single-level quantum dot placed between two metallic leads held at different temperatures and voltages.⁶ Positioning the dot level far enough from the Fermi levels of the leads enables the electrons to flow against the voltage bias while carrying heat from hot to cold. Thermoelectric properties of weakly coupled quantum dots have also been experimentally studied.^{7–10} This type of device operates as a heat engine by generating electrical current and transferring heat between the same two reservoirs. A recent modification^{11,12} of this scheme makes the charge and heat currents flow along different pathways by introducing a third reservoir: charge is transported between two reservoirs at the same temperature, while the third reservoir, at a different temperature, is Coulomb-coupled to the transport electrons and supplies the thermal fluctuations driving the heat engine.

Here we introduce a new type of mesoscopic heat engine, sketched in Fig. 1(a). It consists of one or more quantum dots or metallic islands between two electronic leads at the same temperature T but with a voltage bias V . The leads are connected to an external circuit with temperature T_E and impedance $Z(\omega)$. If the tunnel coupling for one of the junctions is small enough, a tunneling electron will exchange energy with the electromagnetic

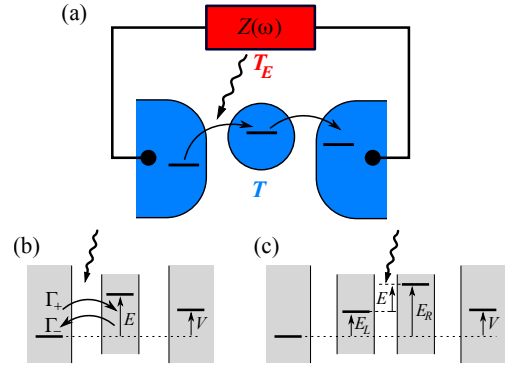


FIG. 1. (color online). (a) Schematic of a single-electron heat engine coupled to an external environment with impedance $Z(\omega)$. Electrons tunneling through the left junction exchange energy with the heat bath, enabling net current against the voltage bias. (b), (c) Energy-level diagrams of junction systems with one or two dots between the leads. One of the junctions, marked with a photon symbol, is coupled to the external bath. The photon-assisted tunneling rates are Γ_{\pm} .

environment, and with a proper choice of parameters it is then possible to achieve a net current against the voltage bias. This system is similar in spirit to the devices of Refs. 11 and 12 but instead of a direct Coulomb interaction between the electrons in the transport system and in the heat bath, in our proposal the coupling is mediated by microwave photons. Photonic heat conduction in electronic circuits has been investigated in recent years both theoretically^{13–18} and experimentally,^{19,20} while previously studied single-electron devices for thermal applications include a thermometer^{2,21}, a cooler¹⁴, and a heat diode.²²

A significant advantage of the photonic coupling is that the two parts of the circuit with different temperatures can be located arbitrarily far apart. One can imagine a scenario where the external circuit is a relatively large device which performs some useful function but at the same time produces excess heat. Our engine can recover a part of this heat and feed it back to the main device as

electrical power.

We show that in an optimal configuration the electrical current in different types of junction systems is given by a single concise formula, Eq. (8). Then a junction between two quantum dots is shown to be ideal in terms of efficiency while maximal power can be provided by a junction between a quantum dot and a metal. The effect of the environment spectrum is seen to be relatively minor.

The rest of the paper is organized as follows. In Sec. II we present the theory for electron tunneling in a linear external environment. In Sec. III we consider bath-assisted transport in junction arrays and conclude that in an appropriate limit the current is given by Eq. (8) for a variety of different systems. The heat engine characteristics of different junction types are studied in detail in Sec. IV, with a particular emphasis on performance under maximum power conditions. In Sec. V we summarize the results and consider experimental prospects.

II. COUPLING TO THE ENVIRONMENT

The exchange of energy between the tunneling electron and the external environment is treated with the so-called $P(E)$ theory. A thorough account of this formalism is given in Refs. 23 and 24, and in this Section we only present the results relevant for the present study. The fundamental assumption is that the coupling between two electron systems is weak enough so that tunneling between them can be adequately described by the Fermi golden rule and that the environment relaxation is much faster than the tunneling rate. Then within $P(E)$ theory the tunneling rate $\Gamma_{i \rightarrow j}$ between the electron systems i and j is given by the Fermi golden rule formula with the energy-conserving δ function replaced by $P(E)$, the probability density to exchange energy E with the environment. Thus we have²³ (with $\hbar = k = e = 1$)

$$\Gamma_{i \rightarrow j} = 2\pi |t|^2 \int d\varepsilon_i d\varepsilon_j \rho_i(\varepsilon_i - \mu_i) \bar{\rho}_j(\varepsilon_j - \mu_j) P(\varepsilon_{ij}) \quad (1)$$

where $\varepsilon_{ij} = \varepsilon_i - \varepsilon_j$. For the electron density $\rho_i(\varepsilon)$ we consider two cases, a single-level quantum dot with $\rho_i(\varepsilon) = \delta(\varepsilon)$, and a metal with $\rho_i(\varepsilon) = \nu_i f(\varepsilon)$ where ν_i is the density of states and $f(\varepsilon)$ is the Fermi function. Similarly the hole density is $\bar{\rho}_i(\varepsilon) = \delta(\varepsilon)$ for dots and $\bar{\rho}_i(\varepsilon) = \nu_i [1 - f(\varepsilon)]$ for metals. Note that in the quantum dot case we assume that dot i is occupied and dot j is empty, otherwise the rate would vanish. The tunneling matrix element between the initial and final states is t , which is taken here to be energy independent. The Fermi level of a metal or the single level of a dot is μ_i with possible Coulomb charging energies absorbed into it. The heat current emitted by the environment during the tunneling process is obtained from the rate formula

by weighing the integral with $\varepsilon_j - \varepsilon_i = -\varepsilon_{ij}$:

$$J_{i \rightarrow j} = -2\pi |t|^2 \int d\varepsilon_i d\varepsilon_j \varepsilon_{ij} \rho_i(\varepsilon_i - \mu_i) \bar{\rho}_j(\varepsilon_j - \mu_j) P(\varepsilon_{ij}) \quad (2)$$

The $P(E)$ function for an electromagnetic environment can be determined by a circuit theory analysis of the system. If the junction, which itself has some capacitance C , is coupled to an environment with impedance $Z(\omega)$, the total impedance Z_t over the junction is C and $Z(\omega)$ in parallel, that is, $Z_t(\omega) = [i\omega C + 1/Z(\omega)]^{-1}$. We will omit the rather complicated general expression for $P(E)$. For our purposes it is sufficient to note that it only depends on the environment temperature T_E and the real part of $Z_t(\omega)$. Since $P(E)$ is a probability density its integral is normalized to unity, and additionally the detailed balance for the environment requires that²³ $P(-E) = e^{-E/T_E} P(E)$.

We will consider two simple and prototypical environments, a harmonic oscillator and an ohmic resistor. The form of the $P(E)$ function for an oscillator with frequency ω_0 can be argued rather intuitively. When an electron tunnels through the capacitance C , the expectation value of the increase in the energy of the electromagnetic field is equal to the charging energy $E_C = e^2/2C$, corresponding to E_C/ω_0 emitted photons at $T_E = 0$. At finite temperatures the electron can also absorb energy from the field, and the expected numbers of absorbed and emitted photons are²³ $\lambda_a = \frac{E_C}{\omega_0} N$ and $\lambda_e = \frac{E_C}{\omega_0} (N + 1)$, respectively, with $N = (e^{\omega_0/T_E} - 1)^{-1}$. The $P(E)$ function can then be written as a double Poisson process for absorbing $m \geq 0$ photons and emitting $m + k \geq 0$ photons, resulting in a net emission of $E = k\omega_0$:²³

$$P(E) = e^{-\lambda_a - \lambda_e} \sum_{\substack{k \in \mathbb{Z} \\ m \geq \max\{0, -k\}}} \frac{\lambda_a^m \lambda_e^{m+k}}{m!(m+k)!} \delta(E - k\omega_0) \quad (3)$$

We consider two special cases of this equation, namely the limits of large and small ω_0 . When $\omega_0 \gg T_E$, we have $N \approx e^{-\frac{\omega_0}{T_E}}$ and for each k the term with $m = \max\{0, -k\}$ dominates the sum while others are exponentially suppressed. This results in

$$P(E) = e^{-\frac{E_C}{\omega_0}} \sum_{k \in \mathbb{Z}} \frac{e^{\frac{k\omega_0}{T_E} \theta(-k)}}{|k|!} \left(\frac{E_C}{\omega_0}\right)^{|k|} \delta(E - k\omega_0) \quad (4)$$

where θ is the unit step function.

In the opposite limit, $\omega_0 \ll T_E$, we have $N \approx \frac{T_E}{\omega_0}$, and the expected photon numbers $\lambda_a \approx E_C T_E / \omega_0^2$ and $\lambda_e = \lambda_a + E_C / \omega_0$ become large so that the two Poisson processes in Eq. (3) can be replaced by normal distributions with mean and variance $\lambda_{a/e}$. The probability to emit a net energy E is a convolution of these two Gaussians, itself also a Gaussian:²³

$$P(E) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(E - E_C)^2}{2\sigma^2}} \quad (5)$$

where $\sigma^2 = 2E_C T_E$. An important fact about this equation is that it also describes a highly resistive ohmic environment, which can be seen by noting that a harmonic oscillator with $Z(\omega) = i\omega L$ has $\text{Re } Z_t(\omega) = \frac{\pi}{2C}[\delta(\omega - \omega_0) + \delta(\omega + \omega_0)]$, while a resistor with $Z(\omega) = R$ has $\text{Re } Z_t(\omega) = R/[1 + (\omega RC)^2]$. Taking the small-frequency limit of the oscillator, $\omega_0 \rightarrow 0$, and the large-impedance limit of the resistor, $R \rightarrow \infty$, will both result in $\text{Re } Z_t(\omega) \rightarrow \frac{\pi}{C}\delta(\omega)$. Therefore both of these systems will also have the same $P(E)$ given by Eq. (5), known as the high-impedance environment. The spectrum of a low-resistance ohmic environment, on the other hand, consists of a broadened delta peak at the origin,²³ approximating vanishing environmental coupling and will therefore not be considered any further. Equations (4) and (5), representing discrete and continuous environments, will be used to investigate the effect of the bath spectrum shape on the performance of the heat engine.

III. CURRENTS IN JUNCTION SYSTEMS

Since the aim of the present paper is to study thermoelectric power generation in tunnel junctions interacting with a heat bath through electromagnetic fields, it is natural to start with the simplest possible system of this type: a single junction between two noninteracting metallic leads coupled to an environment with arbitrary $P(E)$. The tunneling rate for the positive direction, from left (L) to right (R), is Γ_+ and the rate in the opposite direction is Γ_- . With the voltage bias $\mu_R - \mu_L = V$, Eq. (1) gives the current as

$$I = \Gamma_+ - \Gamma_- \quad (6)$$

$$= \gamma \int d\varepsilon_L d\varepsilon_R f(\varepsilon_L)[f(V - \varepsilon_R) - f(-V - \varepsilon_R)]P(\varepsilon)$$

where $\gamma = 2\pi|t|^2\nu_L\nu_R$ and $\varepsilon = \varepsilon_L - \varepsilon_R$. We use a convention where a positive voltage V means that electrons have a higher energy in the right lead and a positive current I means that electrons travel, on the average, from left to right. Therefore if I and V have the same sign, that is, the power $P = IV$ is positive, heat is converted to electrical energy and the device operates as a heat engine. For the present case the bracketed term in Eq. (6) has a sign that is opposite to the sign of V and therefore $P \leq 0$. This simple junction cannot produce thermoelectric power. One way to understand this failure is to think of the system as a Brownian motor²⁵ which is driven by thermal fluctuations due to the external environment. However, these fluctuations do not intrinsically have any preferred direction and therefore the noise must be rectified by the electron transport system in order to generate net power.¹² Rectification requires a nonlinear current-voltage characteristic but the simple junction (without the environment coupling) is linear. Therefore a nonlinearity must be introduced into the transport system, and here we do it by adding a quantum dot or a metallic island between the two leads. We further assume that

the Coulomb repulsion within the central dot is so strong that it can only be empty (with probability P_0) or singly occupied (with probability P_1). The Fermi level of the metallic island or the single-particle level of the quantum dot is at energy E , with Coulomb energy included; see Fig. 1(b). The tunneling rates through the left and right junction are Γ_\pm and $\Gamma_{R\pm}$, respectively. Now, instead of Eq. (6), the current is given by $I = P_0\Gamma_+ - P_1\Gamma_-$. The probabilities can be solved from the master equation $\dot{P}_0 = P_1(\Gamma_- + \Gamma_{R+}) - P_0(\Gamma_+ + \Gamma_{R-})$ in the steady state, $\dot{P}_0 = 0$. This yields the current

$$I = \frac{\Gamma_+\Gamma_{R+} - \Gamma_-\Gamma_{R-}}{\Gamma_+ + \Gamma_{R+} + \Gamma_- + \Gamma_{R-}} \quad (7)$$

If the two junctions are identical, symmetry of the system implies that there can be no thermally induced current for $V = 0$, and a calculation with Eq. (1) shows that there can be no power production for any V . For instance, in the case of a quantum dot the numerator of Eq. (7) is proportional to $\int d\varepsilon d\varepsilon' P(\varepsilon)P(\varepsilon')[f(\varepsilon + E)f(\varepsilon' - E + V) - f(\varepsilon + E - V)f(\varepsilon' - E)]$. For $V > 0$ the first and the second Fermi function are smaller than the third and the fourth, respectively, implying $I < 0$. To arrive at the preceding expression the $P(E)$ functions for the two junctions must be identical, and therefore unequal environmental couplings are required for power production.

Since both junctions are connected to the same external circuit, the only way they can have different $P(E)$ functions is that the junction capacitances, C_L and C_R , are unequal. In that case circuit theory analysis²³ shows that the total impedance over junction i is $\text{Re } Z_{t,i}(\omega) = \kappa_i^2 \text{Re } \tilde{Z}_t(\omega)$, where $\kappa_i = C/C_i$ and $\tilde{Z}_t(\omega) = [i\omega C + 1/Z(\omega)]^{-1}$, with $1/C = 1/C_L + 1/C_R$. In other words, the only difference to a single junction is that the external impedance seen by junction i is effectively reduced by a factor κ_i^2 . Thus a maximal difference between the two $P(E)$ functions can be obtained when the capacitances have different orders of magnitude. If, for example, $C_L \ll C_R$, we have $C \approx C_L$, $\kappa_L \approx 1$, and $\kappa_R \approx C_L/C_R$. In this limit the left junction is coupled to the environment with energy $E_C = e^2/2C_L$ while the coupling of the right junction is suppressed by a factor of $(C_L/C_R)^2$. Note however that the charging energy is $e^2/2C_R$ which must still be large enough to prevent a multiple occupation of the island. To simplify Eq. (7) we note that since the right junction is effectively decoupled from the environment, detailed balance requires $\Gamma_{R+} = \Gamma_{R-}e^{(E-V)/T}$. This relation can, of course, be also derived from Eq. (1) with $P(E) = \delta(E)$. Furthermore, a junction with a small capacitance is usually also weakly transmitting and therefore it is consistent to assume that $C_L \ll C_R$ implies $\Gamma_\pm \ll \Gamma_{R\pm}$. Then we obtain

$$I = f(V - E)\Gamma_+ - f(E - V)\Gamma_- \quad (8)$$

This is the main equation of the present paper. Below we argue that it applies also to more generic junction systems, and its heat engine characteristics are examined in detail in the next section.

Let us then turn to the case of two metal islands or quantum dots in series between the leads, as depicted in Fig. 1(c). We again assume a strong enough Coulomb repulsion so that the double dot can be either empty (probability P_0), or there can be one electron in the left dot (probability P_L) or in the right dot (probability P_R). The tunneling rates through the left, center, and right junction are $\Gamma_{L\pm}$, Γ_{\pm} , and $\Gamma_{R\pm}$, respectively, and the current is given by $I = P_L\Gamma_+ - P_R\Gamma_-$. Master equation for the occupations is

$$\begin{aligned}\dot{P}_0 &= -(\Gamma_{L+} + \Gamma_{R-})P_0 + \Gamma_{L-}P_L + \Gamma_{R+}P_R \\ \dot{P}_L &= \Gamma_{L+}P_0 - (\Gamma_{L-} + \Gamma_+)P_L + \Gamma_-P_R \\ \dot{P}_R &= \Gamma_{R-}P_0 + \Gamma_+P_L - (\Gamma_{R+} + \Gamma_-)P_R\end{aligned}\quad (9)$$

and the steady-state solution gives the current as

$$I = (\Gamma_{L+}\Gamma_+\Gamma_{R+} - \Gamma_{L-}\Gamma_-\Gamma_{R-})/\tilde{\Gamma}^2 \quad (10)$$

with $\tilde{\Gamma}^2 = \Gamma_{L+}\Gamma_{R+} + \Gamma_{L-}\Gamma_{R-} + \Gamma_{L-}\Gamma_{R+} + \Gamma_+(\Gamma_{L+} + \Gamma_{R+} + \Gamma_{R-}) + \Gamma_-(\Gamma_{L+} + \Gamma_{L-} + \Gamma_{R-})$. Similarly to the two-junction case, we now assume that the central junction has a small capacitance and a small transmittance compared to the other junctions. Thus $\Gamma_{\pm} \ll \Gamma_{L\pm, R\pm}$, and since the left and right junctions are effectively decoupled from the environment, detailed balance implies $\Gamma_{L-}/\Gamma_{L+} = e^{E_L/T}$ and $\Gamma_{R+}/\Gamma_{R-} = e^{(E_R-V)/T}$. The expression for the current is then simplified to

$$I = \frac{e^{(E-V)/T}\Gamma_+ - \Gamma_-}{1 + e^{(E-V)/T} + e^{(E_R-V)/T}} \quad (11)$$

The last term of the denominator is vanishingly small compared to the first two if $-(E_R - V) \gg T$ and $-E_L \gg T$. In this limit the expression for the current is again reduced to Eq. (8). Physically this limit means that the dot levels are much below the Fermi levels of the leads and thus either of the two dots is always occupied ($P_0 \rightarrow 0$), leading effectively to a two-state system similar to the one-dot case. Clearly this argument could also be extended to a larger number of dots if necessary.

We have now seen that Eq. (8) is the fundamental expression for thermoelectric current in the optimal limit when only one junction is exchanging heat with the external bath and when all “idle” time spent on processes with no environment coupling can be neglected. The generality of Eq. (8) can also be intuitively seen as follows. The system is always in one of two possible states: there is an electron ready to tunnel either from left to right or from right to left through the bath-coupled junction. The probabilities of these two states are P_+ and P_- , and the average current through the junction is therefore $I = P_+\Gamma_+ - P_-\Gamma_-$. On the left side of the junction there is a metal with Fermi level E_L , or a quantum dot with a level at E_L strongly coupled to a metal. In either case, the probability that the left side has an electron ready for tunneling is proportional to $f(E_L)$. Similarly the right side has a level at $E_R - V$, and it is ready to receive the tunneling electron with a probability proportional to $1 - f(E_R - V) = f(V - E_R)$.

Thus we have $P_+ \propto f(E_L)f(V - E_R)$ and analogously $P_- \propto f(-E_L)f(E_R - V)$. Normalizing with $P_+ + P_- = 1$ gives $P_0 = f(V - E)$ and $P_1 = f(E - V)$, and we arrive at Eq. (8). The existence of these two states is due to the electron-electron interaction. The noninteracting case of Eq. (6) is obtained when $P_+ = P_- = 1$, that is, when there is only one single state allowing electrons to tunnel at any time in either direction.

The heat current J emitted by the environment can be calculated by following exactly the same steps as for the electrical current, and the result corresponding to Eq. (8) is

$$J = f(V - E)J_+ + f(E - V)J_- \quad (12)$$

where J_+ and J_- are the energy currents absorbed by the electron tunneling to the right and to the left, respectively, as obtained from Eq. (2). The sign difference between Eqs. (8) and (12) is due to the fact that J_{\pm} already contain the direction of the heat flow.

IV. HEAT ENGINE CHARACTERISTICS

As we have argued, in the optimal limit an array of one or more quantum dots or metal islands can be seen as a two-state system obeying Eqs. (8) and (12), and thus when evaluating the thermoelectric power generation in these devices it is only necessary to consider the single junction that is coupled to the electromagnetic environment. Some generic remarks about Eq. (8) can be made without any explicit model for the junction. First note that a mirror reflection of a solution with (E, V, I) produces another solution with $(-E, -V, -I)$ and therefore it is sufficient to consider the case $E > 0$. Next, the $I(V)$ curve of Eq. (8) has a simple overall structure. Since the tunneling rates Γ_{\pm} depend only on the level difference E over the junction but not on the voltage bias V , the current depends on V only through the Fermi functions. Thus we see that I approaches asymptotically Γ_+ and $-\Gamma_-$ for large negative and positive values of V , respectively. Since $I(V)$ decreases monotonically, at some point $V = V_0$ the current vanishes. For our purposes the most important fact is that for voltages between 0 and V_0 , I and V have the same sign and the device operates as a heat engine. From Eq. (8) we can solve

$$V_0 = E - T \log \frac{\Gamma_-}{\Gamma_+} \quad (13)$$

All the systems studied here have $\Gamma_- > \Gamma_+$ and therefore E is an upper limit for V_0 while no lower limit exists. We will also see that a hot environment ($T_E > T$) implies $V_0 > 0$ and therefore the heat engine operates with positive V and I while for a cold environment ($T_E < T$) V and I are negative.

An important characteristic of a heat engine is the efficiency $\eta = IV/J_H$, where J_H is the heat current from the hot bath. When $T_E > T$ we have simply $J_H = J$

but when $T_E < T$ the hot bath is the transport system and then $J_H = IV - J$, that is, the heat taken from the electrons is the sum of the produced power IV and the heat $-J$ expelled to the cold environment. Thus we have

$$\eta = \begin{cases} \frac{IV}{J}, & T_E > T \\ \frac{J}{IV-J}, & T_E < T \end{cases} \quad (14)$$

The fundamental upper limit for η is the Carnot efficiency $\eta_C = 1 - T_C/T_H$, where $T_C = \min\{T, T_E\}$ and $T_H = \max\{T, T_E\}$. In this limit transport proceeds reversibly and thus power production is vanishingly small. In practice one often wishes to operate the engine with parameters that maximize the *power* instead of efficiency. As shown by, among others, Curzon and Ahlborn,²⁶ using a particular model for a heat engine it is possible to derive a simple upper bound $\eta_{CA} = 1 - \sqrt{T_C/T_H}$ for the efficiency at maximum power. However, later it has been realized that η_{CA} is not a universal limit,²⁷ and in fact, as we will see below, our engine is also able to exceed it.

Another efficiency measure, widely used in the context of thermoelectric power generation, is the figure of merit ZT . Even though the separation of heat and charge pathways makes the devices covered by our theory quite different from typical thermoelectrics, the usual definition of ZT can still be straightforwardly applied to the present case. In the linear response limit, when V and $\Delta T \equiv T_E - T$ are much smaller than T , we have¹

$$ZT = \frac{\sigma S^2}{\kappa} T \quad (15)$$

where the Seebeck coefficient is $S = \partial V_0 / \partial(\Delta T)$, electrical conductance is $\sigma = \partial I / \partial V$ at $\Delta T = 0$, and thermal conductance is $\kappa = \partial J / \partial(\Delta T)$ at $I = 0$. Generally κ should include all forms of heat transfer between the reservoirs, but since we are not modeling any parasitic flows they are not included in ZT . A central goal in thermoelectrics research is to produce systems with $ZT > 1$.¹

When an electron is transported between the leads of a thermoelectric system it always performs the same amount of work $\pm V$ where the sign depends on the direction of tunneling. However, the amount of heat transferred between the thermal baths can vary from one tunneling event to another, and the spectrum of this energy exchange is determined both by the structure of the junction and the environment. We consider two limiting cases for the junction structure. The first is a fully energy-selective junction, where each tunneling event is accompanied by an exchange of a fixed amount of heat $\pm E$. A physical realization for this is a junction between two quantum dots with sharply defined energy levels separated by E . The other extreme is a totally unfiltered junction where essentially any amount of heat can be exchanged. The two physical examples that we consider are a junction between a quantum dot and a metal, and a junction between two metals.

A. Energy-selective junctions

A junction between two quantum dots can only exchange a fixed energy E with the external bath during the tunneling events. Using Eq. (1) together with detailed balance for the environment yields

$$\begin{aligned} \Gamma_+ &= \Gamma_- e^{-E/T_E} \\ \Gamma_- &= 2\pi |t|^2 P(E) \end{aligned} \quad (16)$$

Similarly the heat flows from Eq. (2) are $J_+ = E\Gamma_+$ and $J_- = -E\Gamma_-$, which imply the simple relation $J = EI$. Therefore at all temperatures the ratio of produced power and transferred heat is a constant V/E , a situation known as strong coupling between particle and heat flows. Carnot efficiency can only be achieved by such strongly-coupled systems,^{5,28,29} and in the present case this can be confirmed by noting that the stopping voltage from Eq. (13) is

$$V_0 = E(1 - T/T_E) \quad (17)$$

and substituting $V = V_0$ in Eq. (14) yields $\eta = \eta_C$. Thus Carnot efficiency implies vanishing current and power.

In the linear response regime, that is, to first order in V and ΔT , the current from Eqs. (8) and (16) is

$$I = \frac{f(E)\Gamma_-}{T^2} (E\Delta T - TV) \quad (18)$$

The value of ZT can be inferred directly by noting that since κ in Eq. (15) is evaluated at $I = 0$, the strong-coupling condition implies $J = \kappa = 0$ and therefore $ZT = \infty$, independent of any parameters. This is another indication that energy-selective junctions can be used to construct maximally efficient thermoelectrics.

In order to investigate the physics beyond linear response and to find out the conditions that maximize the power instead of efficiency, we perform a numerical calculation and for that purpose an explicit expression for the environment spectrum is needed. Since an energy-selective junction interacts with the environment only at a single energy E , the full form of the $P(E)$ function does not generally have much significance. There is one caveat, however: Fermi golden rule, which $P(E)$ theory is based on, is not able to treat transitions between discrete states, and therefore our approach fails if the environment and both sides of the junction are discrete. Thus for a junction between two quantum dots we only consider the continuum environment, Eq. (5). Note, however, that even in this case the limit $\sigma \rightarrow 0$ gives a delta function which is beyond $P(E)$ theory.

For each pair of temperatures T and T_E , the generated power $P = IV$ is numerically maximized with respect to the bias V , the level difference E , and the coupling energy E_C , and the results are presented in Fig. 2(a) as a function of $\Delta T/T_0$, where $T_0 = \frac{1}{2}(T + T_E)$ is the average temperature. Beyond linear response the behavior of the system becomes non-symmetric in ΔT . The limiting cases of large temperature differences can, however,

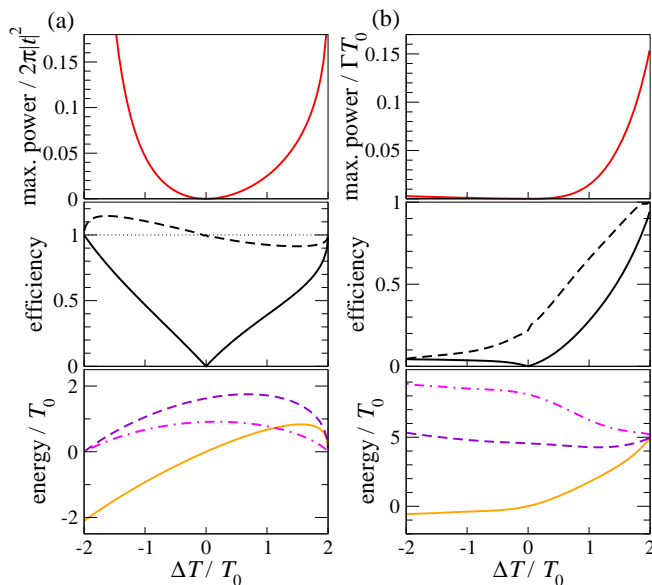


FIG. 2. (color online). Maximum power production for (a) a junction between two quantum dots coupled to a continuum environment, and (b) a junction between a quantum dot and a metal coupled to a discrete environment with $\omega_0 = 5T_0$. Top panels: The maximum power for given temperatures. Note the different units for (a) and (b). Middle panels: efficiency η (full line) and the efficiency compared to the Curzon–Ahlborn value, η/η_{CA} (dashed), at maximum power. Lower panels: values of V (full line), E (dashed), and E_C (dash-dotted) that give the maximum power. The temperature difference is $\Delta T = T_E - T$ and the average temperature is $T_0 = \frac{1}{2}(T + T_E)$.

be understood from simple analytical arguments. When $T \rightarrow 0$ ($\Delta T \rightarrow 2T_0$), we have $V > 0$ and therefore the current must also be positive. In this limit the Fermi functions in Eq. (8) become step functions and we must have $E - V > 0$ so that the forward tunneling process remains while backward tunneling vanishes. On the other hand, to maximize power V should be as large as possible, and therefore we must have $V \rightarrow E$. This is also the condition for maximum efficiency, since in this case all the heat E taken from the hot environment is used for work V , implying $\eta \rightarrow 1$. This does not violate the second law since a vanishingly small energy $E - V$ emitted to a cold bath with $T \rightarrow 0$ is enough for the entropy change $\Delta S = -E/T_E + (E - V)/T$ to be positive. Thus when the cold bath is approaching zero temperature it is possible to have efficiencies close to the Carnot value with finite power. Finally, the actual value of the maximum power and the values of E_C and $E = V$ are determined by the form of $P(E)$. The calculation shows that the optimal case for Eq. (5) has $E_C \rightarrow 0$ and therefore $\sigma \rightarrow 0$. As discussed above, the environment spectrum now becomes a delta peak which is beyond the present theory, leading to an unphysical divergence in the power. This problem only occurs when the two temperatures have very different magnitudes and therefore it is of no practical consequence.

In the opposite limit, $T_E \rightarrow 0$ ($\Delta T \rightarrow -2T_0$), we have $V < 0$ and the current must also be negative. The requirement that $E > 0$ implies $\Gamma_+/\Gamma_- = e^{-E/T_E} \rightarrow 0$ so that tunneling in the positive direction is suppressed. This is because the energy-selectivity of the junction forces the electron to absorb an energy E from the environment during this process, but in the limit $T_E \rightarrow 0$ this energy is not available. From the second Fermi function of Eq. (8) we then see that the magnitude of the current, and thus the power, is maximized when E is as small as possible, and therefore the condition for maximum power is $E \rightarrow 0$. This also maximizes the efficiency since only a vanishingly small energy E needs to be expelled into the cold environment so that in order to do work $|V|$, only the same amount of energy needs to be taken from the hot electron system, again resulting in $\eta \rightarrow 1$. The remaining parameter values are determined by numerical calculations but just as in the previous case, also here the maximization with Eq. (5) results in $\sigma \rightarrow 0$ and in the unphysical divergence of power.

One more indication of the highly efficient nature of the energy-selective junction is the fact that the value of η at maximum power stays rather close to the Curzon–Ahlborn limit. The efficiency at maximum power has been actively studied in recent years and it has been shown^{27,29} that for a system with strongly coupled heat and particle flows the efficiency will approach the Curzon–Ahlborn value in the limit $\Delta T \rightarrow 0$. Such general results are not known for finite temperature differences, and as can be seen from Fig. 2(a), the present system actually has $\eta > \eta_{CA}$ for a cold environment while η_{CA} is an upper bound in the presence of a hot environment.

In addition to the different measures of efficiency, a very important characteristic is the absolute magnitude of power achievable with the device. The unit of power in Fig. 2(a) is $2\pi|t|^2$, and to find an upper limit for this quantity we note that the junction between two quantum dots can be seen as a two-state system with level difference E and tunnel coupling t . The left and right dots must be approximate energy eigenstates, requiring that $|t| \ll E$. Since $E \approx T_0$ for maximum power operation, we end up with $|t| \ll T_0$. Thus the output power is strictly limited by the operating temperature.

Finally we remark that this energy-selective system is very similar to the one studied in Ref. 11. The physical implementations are rather different, with the system of Ref. 11 requiring a total of four quantum dots while for our device two dots are sufficient, but the structures of the energy transfer processes are fundamentally identical, leading to very similar performance figures.

B. Unfiltered junctions

Junctions which do not place any restrictions on the amount of energy exchanged between the tunneling electron and the environment will be called *unfiltered*. We

mostly concentrate on a system where the other side of the junction is a quantum dot and the other side is a metal, in which case Eq. (1) gives

$$\Gamma_{\pm} = \Gamma \int d\varepsilon P(\varepsilon) f(\varepsilon \pm E) \quad (19)$$

where $\Gamma = 2\pi|t|^2\nu$. This form shows immediately that $\Gamma_- > \Gamma_+$ and thus $V_0 < E$. Linearizing Eqs. (8) and (12) then yields

$$I = \frac{f(E)e^{\frac{E}{T}}\Gamma}{T^2}(\tilde{E}\Delta T - cTV) \quad (20)$$

$$J = \frac{f(E)e^{\frac{E}{T}}\Gamma}{T^2}(\tilde{J}\Delta T - \tilde{E}TV) \quad (21)$$

while the corresponding figure of merit from Eq. (15) is

$$ZT = \left(\frac{c\tilde{J}}{\tilde{E}^2} - 1 \right)^{-1} \quad (22)$$

where $c = F[0]$, $\tilde{E} = F[1]$, $\tilde{J} = F[2]$, and $F[n] = \int_0^\infty d\varepsilon P(\varepsilon)\varepsilon^n[f(E-\varepsilon)e^{-\frac{\varepsilon}{T}} + (-1)^n f(E+\varepsilon)]$. These definitions show that if $P(E)$ consists of a pair of delta peaks at energies $\pm E_0$, then in the limit $E = E_0 \gg T$ we have $ZT \rightarrow \infty$ and therefore this kind of environment is able to mimic the effect of an energy-selective junction. However, in this limit \tilde{E} , and thus the power IV , scales as $\exp(-E/T)$. A more detailed analysis shows that beyond $ZT \sim 1$ a linear increase in ZT corresponds to an exponential suppression of generated power. This should be contrasted to the energy-selective junction where the infinite ZT is due to the idealized assumption of perfectly sharp energy levels. In reality the levels are broadened and the figure of merit is finite. However, when the level width is decreased, power is reduced roughly inversely, and not exponentially, with ZT .³⁰ Thus an energy-selective junction is the only realistic way of achieving very high efficiencies, at least in the linear regime. The fundamental difference between the energy filtering provided by the junction and by the environment is the fact the former is directional while the latter is not: when an electron tunnels to the right in an energy-selective junction it must always absorb a photon and when tunneling in the other direction it must emit a photon, but for an unfiltered junction both emission and absorption are possible for either direction.

After concluding that large efficiencies are not available for linear response, we turn to the nonlinear regime and conditions for maximum power. Figure 2(b) shows the results for a discrete environment as represented by Eq. (4) with $\omega_0 = 5T_0$. The figure of merit at maximum power is $ZT \approx 0.5$. In the limit of large temperature differences we see that the case of a hot environment, $T \rightarrow 0$ or $\Delta T \rightarrow 2T_0$, is very similar to the energy-selective junction. Indeed the physics is essentially the same: when $V \rightarrow E = E_C = \omega_0$, only the forward process of Eq. (8) remains, always absorbing a quantum ω_0 from

the hot environment and performing the same amount of work, leading to $\eta \rightarrow 1$. In the opposite case of a cold environment, $T_E \rightarrow 0$, performance of the heat engine is dramatically degraded. Since now $V < 0$, electrons tunneling to the left perform useful work, and for the energy-selective system we were able to have high efficiency and power because Γ_+ was vanishing. But a detailed inspection of Eqs. (4) and (19) reveals that now Γ_+ remains large because electrons are able to tunnel to the right by using the thermal energy of the electron system, without absorbing anything from the external bath. The transport processes which are decoupled from the environment produce a considerable leakage current down the voltage bias, thus making heat engine performance very poor.

For practical purposes an ohmic environment is an important special case. This is modeled with the high-impedance $P(E)$ function from Eq. (5), which, as we have noted, also represents the single-mode environment in the limit $\omega_0 \rightarrow 0$. Numerical results for power maximization are shown in Fig. 3(a). In the linear regime the figure of merit is $ZT \approx 0.5$. For $T_E > T$ the maximum power curve is very similar to the case of discrete environment, and indeed a calculation shows that for $\omega_0 \lesssim 5T_0$ the maximum power varies only weakly with ω_0 while for higher resonance energies it is exponentially suppressed in ω_0/T_E . The efficiency is smaller for the continuum environment since the power is generated by elementary tunneling processes each performing the same work V but absorbing a different amount of heat from the hot bath. These elementary processes thus have different efficiencies, each limited by the second law, and the overall average efficiency must therefore be lower than the Carnot value.

The case of a cold environment, $T_E < T$, is somewhat surprising since both power and efficiency are higher than for a discrete environment. The reason for the very poor performance of the discrete case was seen to be the tunneling processes without environmental coupling. But for the high-impedance $P(E)$, with $T_E \ll E_C$, the environment spectrum is sharply peaked about E_C , away from the origin, and the decoupled processes are suppressed. Consequently the maximum power in the $T_E \rightarrow 0$ limit is increased by an order of magnitude.

As seen from Figs. 2(b) and 3(a), for a junction between a metal and a quantum dot the magnitude of the maximum power depends on the product of Γ and T_0 . Since we are assuming that only a single level of the dot is accessible, both of these quantities must be small compared to the dot level spacing but Γ is not directly constrained by T_0 . Therefore in a system with a level spacing much larger than the temperature it is possible to have $\Gamma T_0 \gg T_0^2$.

Finally we consider another model system for the unfiltered junction, namely tunneling between two metals. In this case Eq. (1) yields

$$\Gamma_{\pm} = \gamma \int d\varepsilon P(\varepsilon)\varepsilon n(\varepsilon \pm E) \quad (23)$$

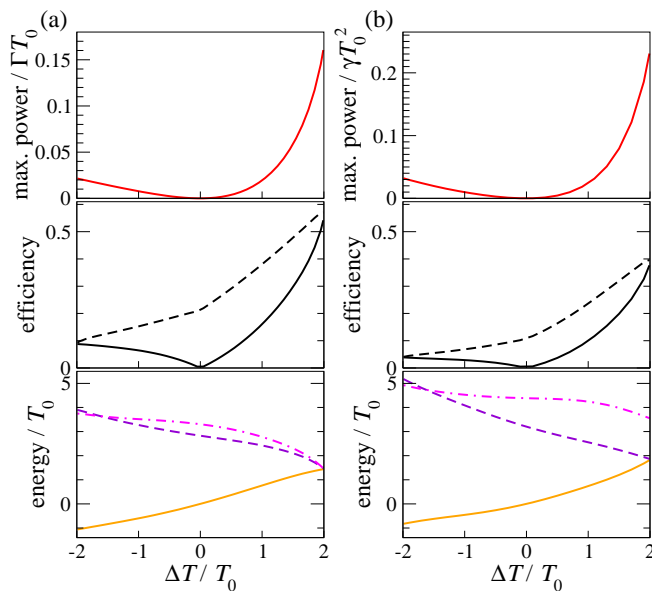


FIG. 3. (color online). Maximum power production for (a) a junction between a quantum dot and a metal coupled to a continuum environment, and (b) a junction between two metals coupled to a continuum environment. See Fig. 2 for an explanation of the different plots.

where $\gamma = 2\pi|t|^2\nu_L\nu_R$ and $n(\varepsilon)$ is the Bose function at temperature T . Comparing this to Eq. (19) we see that the quantitative results for the junction between a quantum dot and a metal can be transferred to the present case by replacing $\Gamma \rightarrow \gamma$ and $f(\varepsilon) \rightarrow \varepsilon n(\varepsilon)$, while all qualitative arguments remain unchanged. We only present the maximum power graph for an ohmic environment in Fig. 3(b) and note the considerable similarity to Fig. 3(a). The efficiency is somewhat lower for the present case, with $ZT \approx 0.2$. The unit of power in Fig. 3(b) is γT_0^2 , and the existence of Coulomb blockade requires²⁴ $\gamma \ll 1$. Therefore just like in the case of the junction between two quantum dots, the operating temperature strictly limits the attainable power.

V. DISCUSSION

In the previous section we have shown that highly efficient energy conversion is in practice only available for an energy-selective junction. This type of device also performs equally well for hot and cold environments. Unfiltered junctions should preferably be operated with a hot environment and cold electron system, and in this case a discrete environment can offer a slightly higher efficiency. If, however, this system is used with a cold external bath, a continuum environment offers better performance.

To intuitively understand the behavior of single-electron heat engines at large temperature differences, one can consider the system in Fig. 1(b) with the left junction coupled to the environment. When the electron

system is cold, that is, T is small compared to the other energies, then the Fermi functions are sharp and electrons cannot tunnel up in energy from the right lead to the center, and therefore only the positive direction transport processes remain. First an electron tunnels from the left lead to the center by absorbing a photon, and then due to the strong tunnel-coupling of the second junction it discharges to the right lead. Thus each tunneling electron transfers heat from hot to cold and performs useful work, leading to optimal thermoelectric performance. On the other hand, for a hot electron system the Fermi functions are smeared and in general electrons can tunnel in both directions without absorbing or emitting photons. Such environment-decoupled processes transport electrical current down the voltage bias, that is, they produce Joule heating from work, hence severely degrading the engine power and efficiency. Only a fully energy-selective junction is able to force the tunneling electrons to always interact with the environment, resulting in highly efficient energy conversion.

Typically achieving high efficiency is a major goal in heat engine design but for the purposes of waste heat recovery the thermal input energy of the device can be considered free and abundant, making efficiency an irrelevant quantity. In this case one should instead concentrate on maximizing the power output, as we have done in Figs. 2 and 3. We showed that for a junction between two quantum dots or between two metals the output power is limited by the temperature in such a way that the units of power in Figs. 2(a) and 3(b) must be much smaller than T_0^2 . The same kind of limitation does not apply for a junction between a metal and quantum dot and thus for fixed temperatures this system has the potential for largest power generation. The reason is that for a dot-metal junction the coupling strength is limited by the dot level spacing which can be much larger than T_0 , while the coupling in a dot-dot junction is limited by the level difference E which cannot be much larger than the temperature or current becomes suppressed. In a metal-metal junction the coupling must be much below the conductance quantum or otherwise Coulomb blockade is wiped out, leading to a linear junction with no thermoelectric effect. If $\Gamma \sim T_0 = 1$ K, the maximum power for all three junction types falls in the femtowatt range. This is a typical figure for low-temperature single-electron devices.^{22,30}

In principle it is possible to increase the power production by having several engine units in parallel with the external impedance. The coupling energy E_C scales inversely with the number of parallel devices and therefore the capacitances of the individual junctions should be decreased. However, in order to see an actual increase in current and power, this change in the capacitances should not considerably lower the tunneling rates of the junctions.

Even if one is not concerned with efficiency, heat leaks between the thermal baths should be minimized in order to have a maximally large temperature difference.

Phononic heat conduction is a problem with all thermoelectric systems. Since phonons are not part of our model we only note that at very low temperatures the electron-phonon coupling becomes weak and this conduction channel can be ignored. On the other hand, two heat leak mechanisms particular to the type of devices considered here are heat conduction by electrons moving between the junction and the external circuit, and photonic heat transfer between the external impedance and the metallic reservoirs of the junction system. The first leakage channel can be eliminated by using superconducting wires, and the second one is suppressed if the resistance of the electron reservoirs is much smaller than the tunnel resistance of the junctions.

One should also note that since the generated current flows through the external impedance $Z(\omega)$, the power P and voltage bias V must be related by $Z(\omega \rightarrow 0) = V^2/P$ if there are no other voltage sources. For example, from Fig. 3(b) we then have $Z(0) \sim 10^2 \gamma^{-1}$ with $\gamma \ll 1$ so that the external resistance must be several orders of magnitude larger than the resistance quantum, which is consistent with the assumption of a high-impedance environment. Of course a non-thermoelectric voltage source

in series with the heat engine can be used to drive the current through an even larger load.

In conclusion, we have investigated the heat engine performance of several types of single-electron junctions coupled to linear electromagnetic environments. Highest thermoelectric performance is obtained when only a single junction is coupled to the external heat bath, and in this case a simple formula is able to describe the essential dynamics. It was confirmed that an energy-selective junction between two quantum dots is capable of highly efficient energy conversion, as evidenced by the high ZT number and the efficiency at maximum power being close to the Curzon-Ahlborn value. However, a junction between two metals can reach the same magnitude of maximum power, while highest power generation is available with a junction between a metal and a quantum dot.

ACKNOWLEDGMENTS

One of the authors (T.O.) would like to thank the Academy of Finland for support.

-
- ¹ A. Shakouri, Annu. Rev. Mater. Res. **41**, 399 (2011).
 - ² F. Giazotto, T. T. Heikkilä, A. Luukanen, A. M. Savin, and J. P. Pekola, Rev. Mod. Phys. **78**, 217 (2006).
 - ³ G. D. Mahan and J. O. Sofo, Proc. Natl. Acad. Sci. USA **93**, 7436 (1996).
 - ⁴ M. Campisi, P. Hänggi, and P. Talkner, Rev. Mod. Phys. **83**, 771 (2011).
 - ⁵ N. Brunner, N. Linden, S. Popescu, and P. Skrzypczyk, arXiv:1106.2138.
 - ⁶ T. E. Humphrey, R. Newbury, R. P. Taylor, and H. Linke, Phys. Rev. Lett. **89**, 116801 (2002).
 - ⁷ A. A. M. Staring, L. W. Molenkamp, B. W. Alphenaar, H. van Houten, O. J. A. Buyk, M. A. A. Mabeoone, C. W. J. Beenakker, and C. T. Foxon, Europhys. Lett. **22**, 57 (1993).
 - ⁸ A. S. Dzurak, C. G. Smith, C. H. W. Barnes, M. Pepper, L. Martin-Moreno, C. T. Liang, D. A. Ritchie, and G. A. C. Jones, Phys. Rev. B **55**, 10197 (1997).
 - ⁹ R. Scheibner, E. G. Novik, T. Borzenko, M. König, D. Reuter, A. D. Wieck, H. Buhmann, and L. W. Molenkamp, Phys. Rev. B **75**, 041301 (2007).
 - ¹⁰ S. Fahlvik Svensson, A. I. Persson, E. A. Hoffmann, N. Nakpathomkun, H. A. Nilsson, H. Q. Xu, L. Samuelson, and H. Linke, arXiv:1110.0352.
 - ¹¹ R. Sánchez and M. Büttiker, Phys. Rev. B **83**, 085428 (2011).
 - ¹² B. Sothmann, R. Sánchez, A. N. Jordan, and M. Büttiker, arXiv:1201.2796.
 - ¹³ D. R. Schmidt, R. J. Schoelkopf, and A. N. Cleland, Phys. Rev. Lett. **93**, 045901 (2004).
 - ¹⁴ J. P. Pekola and F. W. J. Hekking, Phys. Rev. Lett. **98**, 210604 (2007); J. T. Peltonen, M. Helle, A. V. Timofeev, P. Solinas, F. W. J. Hekking, and J. P. Pekola, Phys. Rev. B **84**, 144505 (2011).
 - ¹⁵ T. Ojanen and T. T. Heikkilä, Phys. Rev. B **76**, 073414 (2007).
 - ¹⁶ T. Ojanen and A.-P. Jauho, Phys. Rev. Lett. **100**, 155902 (2008).
 - ¹⁷ T. Ruokola, T. Ojanen, and A.-P. Jauho, Phys. Rev. B **79**, 144306 (2009).
 - ¹⁸ L. M. A. Pascal, H. Courtois, and F. W. J. Hekking, Phys. Rev. B **83**, 125113 (2011).
 - ¹⁹ M. Meschke, W. Guichard, and J. P. Pekola, Nature **444**, 187 (2006).
 - ²⁰ A. V. Timofeev, M. Helle, M. Meschke, M. Möttönen, and J. P. Pekola, Phys. Rev. Lett. **102**, 200801 (2009).
 - ²¹ J. P. Pekola, K. P. Hirvi, J. P. Kauppinen, and M. A. Paalanen, Phys. Rev. Lett. **73**, 2903 (1994).
 - ²² T. Ruokola and T. Ojanen, Phys. Rev. B **83**, 241404 (2011).
 - ²³ G. L. Ingold and Yu. V. Nazarov, in *Single Charge Tunneling*, edited by H. Grabert and M. H. Devoret, NATO ASI Series B Vol. 294 (Plenum Press, New York, 1992), pp. 21–107.
 - ²⁴ Yu. V. Nazarov and Ya. M. Blanter, *Quantum Transport* (Cambridge University Press, Cambridge, 2009).
 - ²⁵ P. Reimann, Phys. Rep. **361**, 57 (2002).
 - ²⁶ F. Curzon and B. Ahlborn, Am. J. Phys. **43**, 22 (1975).
 - ²⁷ M. Esposito, K. Lindenberg, and C. Van den Broeck, Phys. Rev. Lett. **102**, 130602 (2009).
 - ²⁸ O. Kedem and S. R. Caplan, Trans. Faraday Soc. **61**, 1897 (1965).
 - ²⁹ C. Van den Broeck, Phys. Rev. Lett. **95**, 190602 (2005).
 - ³⁰ N. Nakpathomkun, H. Q. Xu, and H. Linke, Phys. Rev. B **82**, 235428 (2010).